Amit Kumar Singh **Yadav**

MACHINE LEARNING · GENERATIVE AI · COMPUTER VISION · SPEECH RECOGNITION

(765) 7147056 | yadav48@purdue.edu | sites.google.com/view/amit-yadav | amit-kumar-singh-yadav

Summary

A Ph.D. candidate in machine learning with 2 years of industrial work experience in software development, and deploying end-to-end machine learning, computer vision, and speech recognition tools/APIs. In past, I have worked on and lead interdisciplinary projects in India (5+ years), Japan (2 months), Canada (3 months), and USA (2+ years).

Education

2021-Present **Ph.D. in ECE, Purdue University**, advised by Prof. Edward J. Delp ***(straight A/A+)[transcript]***

2016-2020 **B.Tech. in EE with Minors in CSE, IIT Gandhinagar**, early graduated ***(Director's Medal)*** [***certificate***]

Work Experiences

May 2024-Aug 2024 **Research Scientist Intern, Meta AI**, ***(other offer- Apple, Dolby, HP, Samsung)***

- researching with LlaMA Speech team on SpeechLLMs for automatic speech recognition

Aug 2021-Present **Research Assistant, VIPER Lab, Purdue University**, ***(7 first authored and total 15 papers)***

- learnt and published generative methods such as Diffusion models for video generation
- created deep learning methods for classification and localization of deepfake speech/video, and attributing the tools used to generate them (Skills: Python, PyTorch, Docker, Git)[***DARPA SemaFor Program***]
- worked on self-supervised transformer and graph neural networks for deepfake speech/video detection using information from AAC/MP4 compressed bit stream (presented at ICASSP 2023, CVPR-W 2023)
- developed multi-domain transformer neural network for detecting deepfake speech (ICASSP 2024)
- did first study quantifying age, gender, and accent bias in deepfake speech detectors (accepted CVPR-W 2024)

Jan 2020-Aug 2020 **Research Intern, Rakuten**, ***(received full-time Applied Researcher position in Vision team)***

- developed machine learning methods for cancer detection, nuclei segmentation, and extracting information from images of outfit size-table (Skills: Python, PyTorch, TensorFlow, Docker, REST APIs, Git)
- deployed ML outfit recommendation system, in a month, increased traffic by 5% on company's e-commerce website by recommending cross-domain users similar product
- in a team of 2, ideated, build and presented demo of a new product PathoAI: an interactive VR based AI assistant for tumour visualization to CEO India, the team received funds to present to executive team

Aug 2020-Aug2021 **Embedded Software Engineer, Enphase Energy**, ***(received excellence award - top 3 globally)***

- developed software to calibrate 3-phase energy metering chips and APIs for cloud-gateway data transfer
- 3-phase metering chip calibration enabled new products for 3-phase high current commercial spaces, and APIs provided cloud control of metering chips - making it fast and cost efficient to solve 15% metering issues at remote sites (Skills: C, C++, Bash, REST APIs, Git, Protobuf)

May 2019-Aug 2019 **Research Intern, Carleton University, Canada**, ***(devised novel interaction)*** ***[publication]***

- learned fabrication tools, and with mobile interaction experts analyzed and solved issues with existing computer vision based back-of-device mobile interactions (Skills: C++, 3D Printing, User-Study, HCI)
- created and user-studied novel back-of-device mobile interactions using off-the self sensors, outcomes show promising potential of using tangible sensors to improve mobile interaction

May 2018-Jun 2018 **Research Intern, JAIST, Japan**, ***(did quantum calculations on Supercomputer)*** ***[publication]***

- performed quantum mechanical calculations for Disiloxane to theoretically predict its chemical properties
- wrote bash scripts for distributed computation on Cray XC40 supercomputer to find linearization barrier of disiloxane, a chemical of extreme importance in cosmetics (Skills: Bash, Parallel Calculations, Monte Carlo)

Selected Awards & Honors

2021 **PhD Meissner Fellowship**, for 2 consecutive years from School of ECE, Purdue University *USA*
2020 **Overall Outstanding Performance** in Department, Director's Medal, IIT Gandhinagar *India*
2019 **MITACS Globalink Scholarship,** International Research Program for Canada ($6000) *Canada*
2019 **Student Travel Grant,** ACM Interactive Surfaces and Spaces Conference ($1500) *South Korea*
2019 **DAAD-WISE Programme,** Invited by Digital Media Lab (declined), University of Bremen *Germany*
2018 **JASSO Scholarship,** for being "Frontier Engineers with Global Sense" ($3000) *JAIST, Japan*
2018 **Excellence in leadership skills,** Sri Temasek@IIT Gandhinagar Scholarship ($270) *India*

Selected Services and Positions of Responsibility

2021-Present **Reviewer**, **prominent machine learning, signal processing and vision conferences** *USA*
- served as reviewer for ICASSP 2024, ECCV 2024, CVPR 2023, ICMLA 2023, ACM Multimedia 2023 conferences

2021-Present **Graduate Mentor**, **Purdue Vertically Integrated Projects on Image Processing** *USA*
- mentoring, helping and grading undergraduates projects on image processing and computer vision

2019-2020 **Co-founded CoShop**, **to deliver essentials in India during COVID Portal | News-Cover** *India*
- created app with 2200+ grocery shops to door deliver essentials, impacted 5 cities and 1K+ families

2018-2019 **Represented 1600+ students**, **Student Representative, Institute Senate, IIT Gandhinagar** *India*
- nominated by Deans, worked on policies related to student debt, privacy, academic help, and mental health

2017-2018 **Led a team of 100 + students**, **Secretary of Robotics Club, IIT Gandhinagar** *India*
- recruited and trained 120 members, raised budget for 15+ projects and 10+ skill-refinement workshops

2016-2017 **Raised 35,000+ USD**, **in a 10 member sponsorship team of Inter College Cultural Fest** *India*
- raised funds and signed MoUs from existing and 30+ new sponsors, led marketing campaigns of two sponsors

Relevant Coursework

Mathematics Courses: Differential Equations, Calculus, Linear Algebra, Probability

Core Courses: Random Variable, Digital Signal Processing, Advanced Signal Processing, Embedded and Microcontroller, Digital Image Processing I, Statistical Pattern Recognition and Decision Making, Computational Models and Methods, Natural Language Processing, Deep Learning, Computer Vision, Optimization, Digital Video Systems

Received A in all core courses, A+ in Deep Learning, Computer Vision, and Digital Video Systems (Compression)

Selected Publications

Amit Yadav, Kratika Bhagtani, Davide Salvi, Paolo Bestagini, Edward J. Delp, **FairSSD: Understanding Bias in Synthetic Speech Detectors**, IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshop (**CVPR 2024**) [accepted] [***pre-print***]

Amit Yadav, Kratika Bhagtani, Sriram Baireddy, Paolo Bestagini, Stefano Tubaro, Edward J. Delp, **MDRT: Multi-Domain Synthetic Speech Localization**, IEEE International Conference on Acoustics, Speech and Signal Processing (**ICASSP 2024**) [***poster***] [***paper***]

Amit Yadav, Emily Bartusiak, Ziyue Alan Xiang, Paolo Bestagini, Stefano Tubaro, Edward J. Delp, **ASSD: Synthetic Speech Detection Using AAC Encoding Parameters**, IEEE International Conference on Acoustics, Speech and Signal Processing (**ICASSP 2023**)[***oral paper***] [***slides***] [***video***]

Ziyue Alan Xiang, **Amit Yadav**, Paolo Bestagini, Stefano Tubaro, Edward J. Delp, **MTN: Forensic Analysis of MP4 Video Files Using Graph Neural Networks**, IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshop (**CVPR 2023**) [***pdf***] [***poster***]

Amit Yadav, Kratika Bhagtani, Ziyue Alan Xiang, Paolo Bestagini, Stefano Tubaro, Edward J. Delp, **DSVAE: Interpretable Disentangled Representation for Synthetic Speech Detection**, IEEE International Conference on Machine Learning and Applications (**ICMLA 2023**) [***long oral paper***][***pre-print***] [**poster**]

Kratika Bhagtani, **Amit Yadav**, Ziyue Xiang, Paolo Bestagini, Edward J. Delp, **FGSSAT : Unsupervised Fine-Grain Attribution of Unknown Speech Synthesizers Using Transformer Networks**, IEEE Asilomar Conference on Signals, Systems, and Computers (**Asilomar 2023**)[***paper***]

Alexandre Olivé Pellicer, **Amit Yadav**, Kratika Bhagtani, Ziyue Xiang, Zygmunt Pizlo, Irmina Gradus-Pizlo, and Edward J. Delp, **Synthetic Echocardiograms Generation Using Diffusion Models**, IEEE Southwest Symposium on Image Analysis and Interpretation **(SSIAI 2024)** [***pre-print***] [***demo***]